Interactive Computer-Aided Design for Molecular Docking and Assembly

Susana K. Lai-Yuen[1] and Yuan-Shin Lee[2]

[1]University of South Florida, laiyuen@eng.usf.edu
[2]North Carolina State University, yslee@ncsu.edu

ABSTRACT

This paper presents a computer-aided design system for *molecular docking* and *nanoscale assembly*. A lab-built 5-DOF (degree of freedom) haptic device and the driving computational engine have been developed to provide force-torque feedback to the users for computer-aided molecular design (CAMD). The developed *haptic force-torque feedback* will enable researchers to visualize, touch, manipulate and assemble molecules in a virtual environment. The presented techniques can be used in the computer-aided molecular design to provide the researchers a real-time tool to better understand molecular interactions and to evaluate possible pharmaceutical drugs and nanoscale devices. Computer implementation and illustrative examples are also presented in this paper.

Keywords: Computer-aided design, molecular docking, nanoscale assembly, haptic interface.

1. INTRODUCTION

Nanotechnology has become the new limit in science and technology [14]. It consists in the ability to manipulate atoms and molecules to discover new pharmaceutical drugs and to create nanoscale devices with new molecular arrangements. *Molecular docking* is used for drug design where a *ligand*, which is a relatively small molecule, docks onto a *receptor*, which is usually a much bigger molecule, as shown in Figure 1. In Figure 1, the example ligand is a Biotin containing 16 atoms and the receptor is a Streptavidin with 901 atoms from the Protein Data Bank (PDB) with the PDB ID: 1stp [2]. These molecules were modeled without the hydrogen atoms. The task of *nanoscale assembly* is essentially similar to molecular docking since biological molecules are used as components for nanoscale devices [14]. Research has just begun towards the identification of biological molecules to be used as components for nanoscale devices [3], and the assembly of molecules in a 3D environment [8],[10]. To achieve the molecular docking and nanoscale assembly, it is crucial to provide researchers with a computer-aided design (CAD) system that allows real-time interactive visualization and manipulation of molecules in a virtual environment. These techniques will help to better understand molecular interactions, and to evaluate the design of novel pharmaceutical drugs and nanoscale devices.

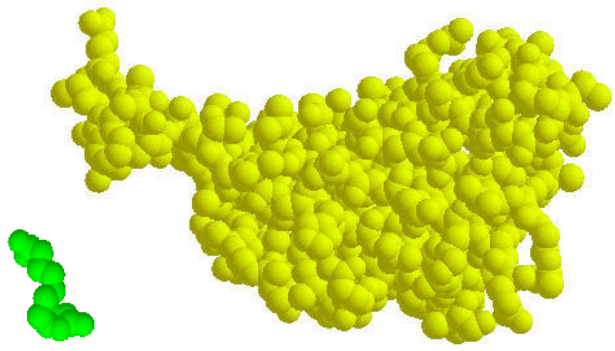


Fig. 1. Receptor and ligand molecules involved in the molecular docking process.

Computer-aided techniques for molecular docking have recently received a lot of attention due to its importance in the medicine and drug design industry while molecular device design is still in its infancy. In recent years, besides using the visualization techniques, there has been increasing interest in using the *haptic interface* to facilitate the exploration and analysis of molecular docking and assembly [10, 12]. A haptic device is an electromechanical device that exerts forces on users giving them the illusion of touching something in the virtual world, as shown in Figure 2.

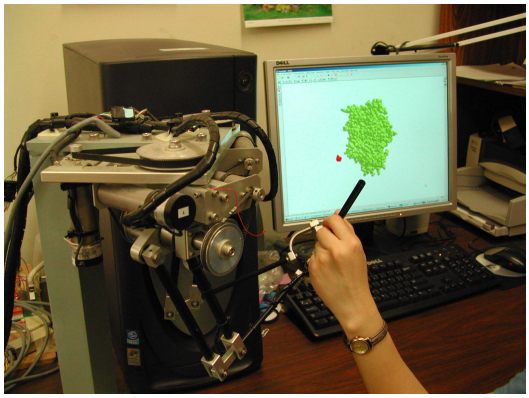


Fig. 2. Haptic device and interface [7].

In our previous work, a haptic-based force-torque feedback system has been developed for product development and manufacturing [13]. To provide the user with realistic sensation of touch through the haptic device, the computation should achieve an update rate of 1 kHz. Previous methods using the haptic interface for molecular docking and assembly modeled the ligand as a rigid body during the analysis and simulation [1],[10],[12]. However, molecules are very "flexible" in adopting different molecular conformations (or shapes) by changing their internal torsional bonds. Therefore, handling molecular flexibility and searching for new and feasible molecular conformations have been identified as computationally expensive and as major challenges in molecular design [9]. Moreover, given the haptic device update rate requirement of 1 kHz, the task of handling molecular flexibility in *real-time* becomes an extremely complex problem.

This paper presents a new method for finding molecular conformations in real-time for molecular docking and assembly. The proposed method will provide researchers with an approximation of the molecule's behavior as molecules are virtually assembled facilitating the identification of pharmaceutical drugs and the design of nanoscale devices.

2. OVERVIEW OF THE PROPOSED MOLECULAR CAD SYSTEM WITH HAPTIC INTERFACE

Figure 3 shows the general concept of the proposed computer-aided design for molecular docking and assembly with haptic interface. The lab-built 5-DOF haptic device, the haptic hardware controller and the haptic rendering software were developed for educational purposes. The haptic interface device can detect 6-DOF of haptic probe movement and provide 5-DOF feedback, with both force and torque feedback.

As shown in Figure 3, the user can manipulate a ligand around the receptor to understand and to study the intermolecular interactions using the developed haptic interface. The user can feel the intermolecular force and torque through the haptic device. In our previous work [4, 7], a nano-scale docking and assembly simulator (NanoDAS) was developed to determine the feasibility in terms of energy and geometry of a ligand to dock or assemble into the receptor. The NanoDAS constructs a search tree starting from the user-specified location to explore the search space and to exploit the low-energy regions. Therefore, as the user manipulates the ligand and feels the intermolecular forces through the haptic device, NanoDAS can help the user in identifying whether a ligand can be docked or assembled into a receptor.

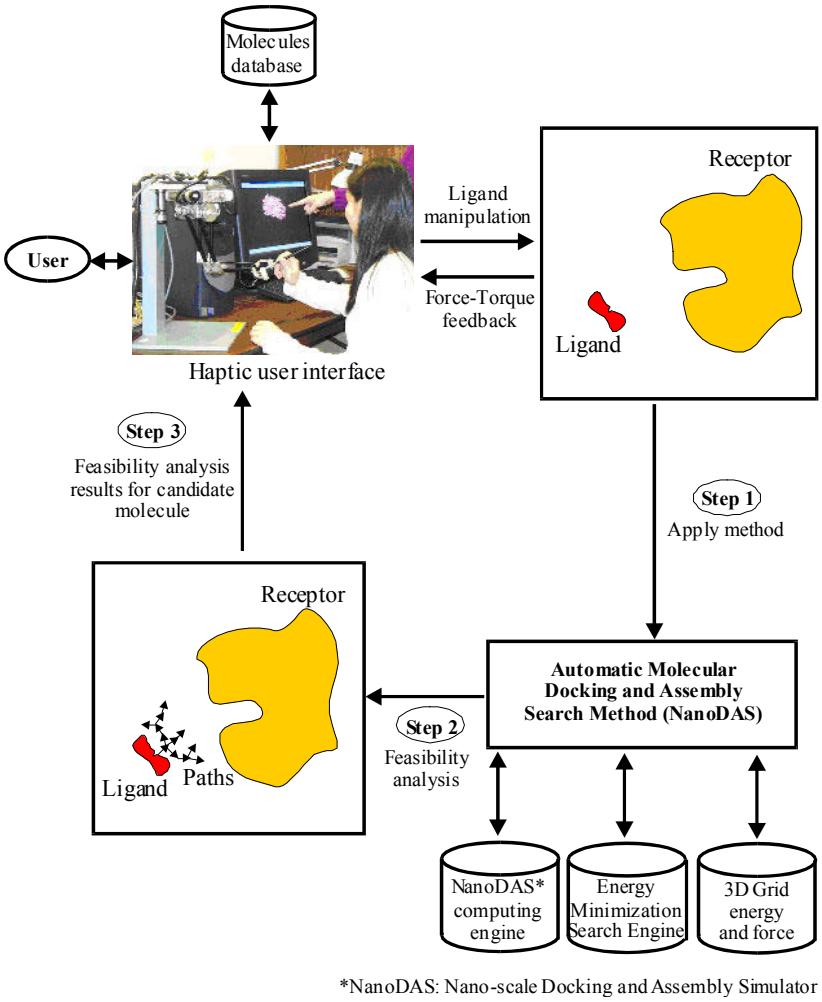


Fig. 3. The proposed haptic-based computer-aided design for molecular docking and assembly.

3. CALCULATION OF POTENTIAL ENERGY AND MOLECULAR INTERACTION FORCE-TORQUE

Interactions between molecules are represented by the potential energy created between them. Although both receptor and ligand molecules are flexible bodies, it is commonly assumed that the receptor is a rigid-body while the ligand is a flexible-body [1]. The potential energy function used in this work is approximated by the van der Waals terms and is expressed as follows:

$$E = \sum_{i=1}^{N_{lig}} \sum_{j=1}^{N_{rec}} \left[\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \right] \tag{1}$$

Where N_{lig} and N_{rec} are the number of atoms in the ligand and the receptor, respectively, A_{ij} and B_{ij} are the van der Waals repulsion and attraction parameters, and r_{ij} is the distance between the centers of atoms i and j. The van der Waals parameters for a single atom type are calculated based on the atom's van der Waals radius and well depth [11]. Given the haptic update requirement of 1 kHz, a 3D grid method is used to calculate the potential energy and force in real-time [1, 10]. The receptor is enclosed by the 3D grid where each grid point stores the potential energy generated by surrounding receptor atoms at that point. A second 3D grid is used to pre-compute the list of possible colliding receptor atoms to speed up collision detection [10]. Each collision grid stores the possible receptor colliding atoms,

assuming a ligand atom is placed in the collision grid cell. When a ligand is introduced in the vicinity of the receptor, all the grids occupied by the ligand atoms are used in calculating the total interaction energy and force between the ligand and the receptor. A 0.5 Å size grid was used in this work since it was shown to provide an accurate enough approximation [1].

The total molecular interaction force $\overrightarrow{F_{total}}$ acting on the ligand by the receptor can be found by summarizing two forces [5, 6]: the molecular potential force $\overrightarrow{F_{potential}}$, and the collision force $\overrightarrow{F_{collision}}$ shown as follows:

$$\overrightarrow{F_{total}} = \overrightarrow{F_{potential}} + \overrightarrow{F_{collision}} = -\frac{dE}{dr_{ij}} + \sum_{i=1}^{n_{collision}} \delta \cdot k \cdot (p_i^{wall} - p_i) \tag{2}$$

Where r_{ij} is the distance between a ligand atom *i* and a receptor atom *j*, $n_{collision}$ is the number of ligand atoms in collision with receptor atoms, δ is the Kronecker delta, and δ=1 if the ligand p_i falls inside any of the collision grids and δ=0 otherwise, p_i is the center of the ligand atom *i* in collision, p_i^{wall} is the projection of p_i onto the virtual wall of the collision receptor, and *k* is a pre-defined spring constant of molecules.

Once the molecular interaction force $\overrightarrow{F_{total}}$ is found by Equation (2), the haptic total torque $\vec{\tau}$ is calculated based on a pivot point P_{pvt} defined as the 'center of weight' of a ligand. The accumulated force is assumed to be applied through the pivot point of the molecule. Therefore, the total torque $\vec{\tau}$ induced at the pivot point P_{pvt} by a collision point P_i is calculated as follows:

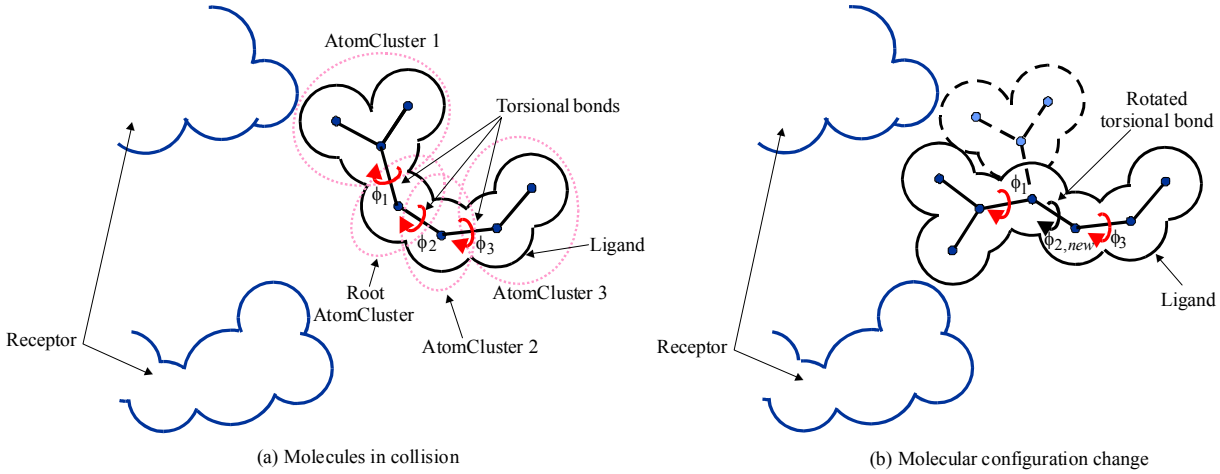
$$\vec{\tau} = \sum_{i=1}^{N_{lig}} \left[\sum_{j=1}^{N_{rec}} \overrightarrow{P_{pvt}P_i} \times \overrightarrow{F_{ij}} \right] \tag{3}$$

Where P_{pvt} is the virtual pivot point's location, P_i is the collision point's location, and $\overrightarrow{F_{ij}}$ is the corresponding collision force between a ligand atom *i* and a receptor atom *j*. The following sections present the haptic force-torque rendering for finding flexible molecular conformations in real-time.

4. HAPTIC INTERFACE FOR REAL-TIME FLEXIBLE LIGAND CONFORMATIONAL SEARCH

As the user manipulates the ligand to experience the molecular interactions with the receptor, the ligand needs to change its geometric shape based on the forces acting on the ligand by the receptor. Anytime the ligand is subject to high forces (or high potential energies), the ligand tends to change its geometric shape or molecular *conformation* to a more stable conformation as shown in Figure 4. A stable molecular conformation is represented by a molecular conformation that minimizes the potential energy *E* described in Equation (1). As shown in Figure 4(a), a molecule can be modeled as an articulated body where an arbitrarily-selected atom acts as the base of the body. A flexible body has at least six degrees of freedom (dof): three translational and three rotational. In addition, each torsional bond of the molecule accounts for an additional degree of freedom, as shown in Figure 4(a). In a molecule, changes in its torsional angles cause the greatest geometric changes compared to changes in the molecule's bond lengths and bond angles. For this reason, it is normally assumed that a molecule's bond lengths and bond angles are fixed during the search for a feasible molecular conformation. Therefore, the conformation of a molecule can be defined as the changes in the angles of the torsional bonds.

Since haptic applications require an update rate of 1 kHz, the computation burden of finding the accurate ligand conformation in real-time becomes infeasible [16]. For this reason, the objective of a quick conformational search is to find a ligand conformation that approximately represents the behavior of a ligand as it approaches the receptor. Figure 4(a) shows a molecule where the atoms are clustered into *AtomClusters* according to the type of bonds in between them [15]. This method creates some local frames to each AtomCluster so the update of atom positions as the torsional bond rotates requires less computation time and decreases inaccuracies in calculations.



(a) Molecules in collision (b) Molecular configuration change

Fig. 4. Molecular conformation change through energy minimization using the adaptive local search method.

To provide the user with real-time ligand molecular flexibility, an adaptive local search is proposed in this paper. Figure 4 shows the process of searching for a new molecular conformation as the ligand collides with the receptor (potential energy increases). When the ligand is determined to be in collision with the receptor as in Figure 4(a), the adaptive local search method is executed to find a set of new torsional angles that result in a new collision-free and low-energy molecular conformation, as shown in Figure 4(b). If no feasible collision-free new molecular conformation can be found after searching, the adaptive local search method returns the original molecular conformation. The following algorithm presents the proposed adaptive local search technique for finding a new collision-free molecular conformation with lower potential energy:

Algorithm. *Adaptive Local Search and Energy Minimization for Molecular Docking*.
Inputs: q_{cur}: Current molecular conformation.
Outputs: q_{new}: New molecular conformation,
E_{new}: Energy for new molecular conformation q_{new} if exists,
$\overrightarrow{F_{new}}$: Haptic force feedback of q_{new},
$\overrightarrow{\tau_{new}}$: Haptic torque feedback of q_{new}.

Step 1. Initialize the new conformation $q_{new} \leftarrow q_{cur}$;
Initialize *ligand_collision* $\leftarrow 0$.

Step 2. Calculate the new molecular energy E_{new} of q_{new} by Equation (1).

Step 3. Calculate the new molecular interactive force $\overrightarrow{F_{new}}$ and the new torque $\overrightarrow{\tau_{new}}$ of the new conformation q_{new} using Equations (2) and (3).

Step 4. Check if any ligand atom of q_{new} occupies any of the collision grids.
If so, set *ligand_collision* $\leftarrow 1$.

Step 5. IF (*ligand_collision* == 0), then
Go to Step 7;
ELSE Set the number of iterations $g \leftarrow 0$.

Step 6. WHILE ($g < g_{max}$)
$q_{test} \leftarrow$ NEW_TORSION_ANGLES();
Set $E_{test} \leftarrow 0$; $\overrightarrow{F_{test}} \leftarrow 0$; $\overrightarrow{\tau_{test}} \leftarrow 0$;
FOR (each AtomCluster AC_k in q_{test})
collision $\leftarrow$ CHECK_ COLLISION(AC_k);
IF (*collision* == 1), Then Exit FOR-Loop;
ELSE, Assign $E_{AC_k}, \overrightarrow{F_{AC_k}}, \overrightarrow{\tau_{AC_k}} \leftarrow$ GET_ENERGY_FORCE_TORQUE(AC_k);

$E_{test} \leftarrow (E_{test} + E_{AC_k})$; $\overrightarrow{F_{test}} \leftarrow (\overrightarrow{F_{test}} + \overrightarrow{F_{AC_k}})$; $\overrightarrow{\tau_{test}} \leftarrow (\overrightarrow{\tau_{test}} + \overrightarrow{\tau_{AC_k}})$.

END-FOR.

IF [(collision == 0) AND ($E_{test} < E_{new}$)]

$E_{new} \leftarrow E_{test}$; $\overrightarrow{F_{new}} \leftarrow \overrightarrow{F_{test}}$; $\overrightarrow{\tau_{new}} \leftarrow \overrightarrow{\tau_{test}}$;

$q_{new} \leftarrow q_{test}$; $g \leftarrow g_{max}$.

END-IF.

$g \leftarrow g + 1$.

END-WHILE.

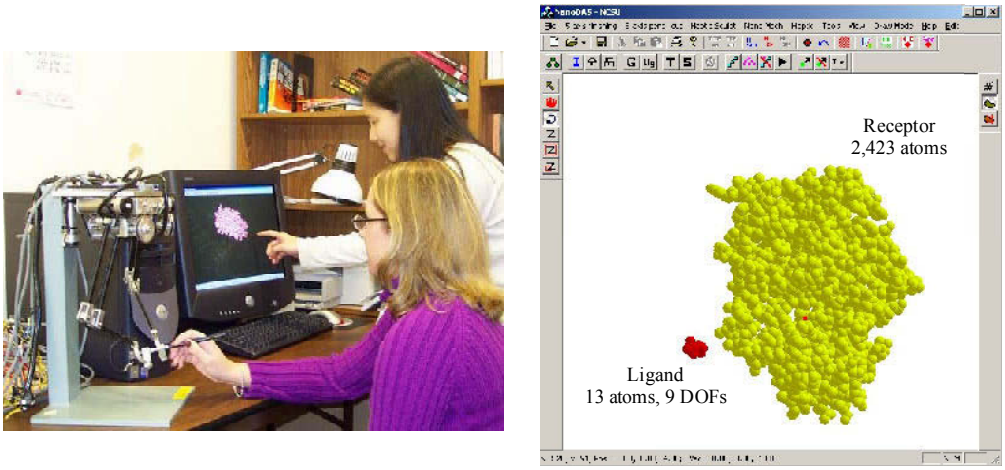
Step 7. Output (q_{new}, E_{new}, $\overrightarrow{F_{new}}$, $\overrightarrow{\tau_{new}}$).

END.

In the algorithm, the function NEW_TORSION_ANGLES() assigns new torsional angles to the ligand within the bonds' allowable torsional angles range for the new test molecular conformation q_{test}. The function CHECK_COLLISION() checks each AtomCluster in the test conformation q_{test} for possible collision with the receptor. The function GET_ENERGY_FORCE_TORQUE() calculates the new energy, interactive force and torque for the new q_{test}. If the test conformation q_{test} is collision-free and its energy E_{test} is below the current conformation's energy E_{new}, the new test conformation q_{test} is accepted to replace q_{new}. More details of the method can be found in [6].

5. COMPUTER IMPLEMENTATIONS AND EXAMPLES

The proposed adaptive searching techniques with the haptic force-torque rendering have been implemented on a dual 2.4 GHz CPU workstation using Visual C++ programming language and OpenGL library functions. The ligands and proteins used in the examples are from the public domain Protein Data Bank (PDB) [2]. Figure 5(a) shows the lab set-up of the haptic device and computer display implemented at our research lab. Figure 5(b) shows the haptic user interface with an example receptor Tyrosyl-tRNA synthetase and a ligand Tyrosine (PDB ID: 4ts1). The Tyrosyl-tRNA synthetase receptor has 2,423 atoms and the Tyrosine ligand has 13 atoms, as shown in Figure 5(b).



(a) Haptic display and device developed at our research lab

(b) Developed haptic interface for molecular docking and assembly

Fig. 5. Haptic graphical interface and the example receptor-ligand complex.

Figure 6 shows the behavior of the ligand as it approaches the example receptor. It can be observed in Figures 6(a) and (b) that the ligand contacts the receptor during docking/assembly and the interaction force is observed. The developed local search algorithm is invoked to find a new feasible conformation of the ligand. Figure 6(c) shows the new conformation of the ligand and the reduced interaction force. As the ligand continues to approach the receptor,

the energy and force increase again as shown in Figure 6(d). Figure 6(e) shows the change of the total potential energy between the ligand and the receptor at different locations. Notice the high energy level at the position *b* where the collision occurs and the lower energy level at the position *c* after it changes to a new flexible conformation, as shown in Figure 6(e). Figure 6(f) shows the interaction force response at different locations as the ligand approaches the receptor using the proposed adaptive local search method is applied.

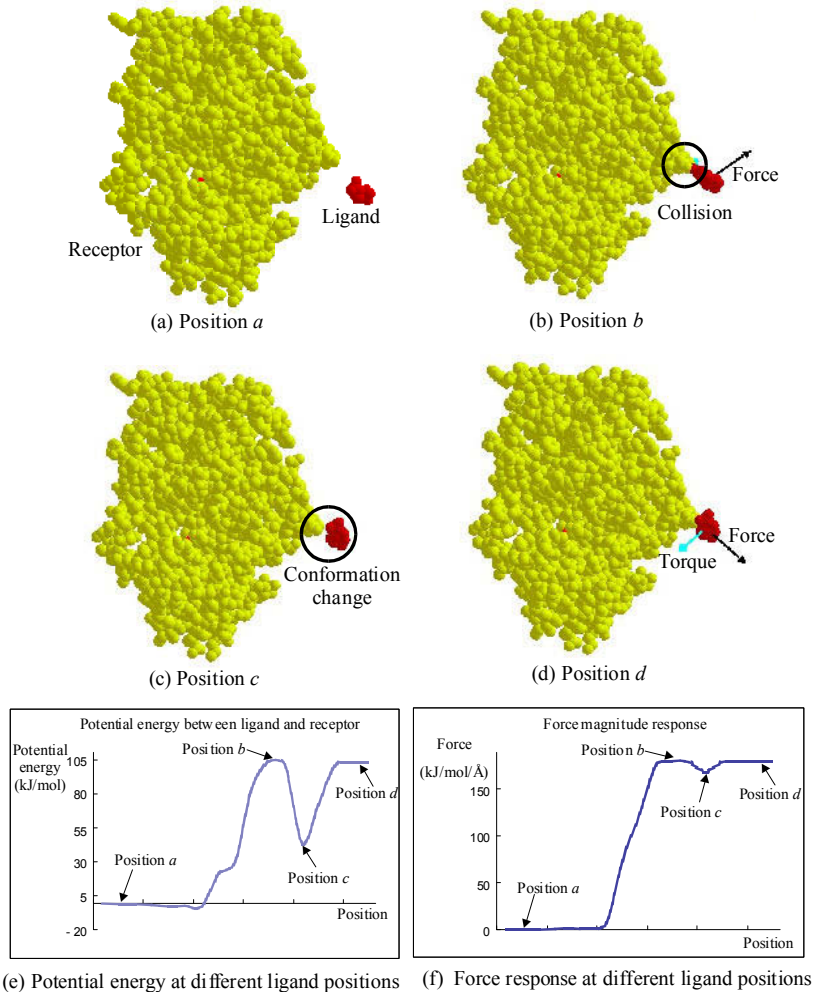
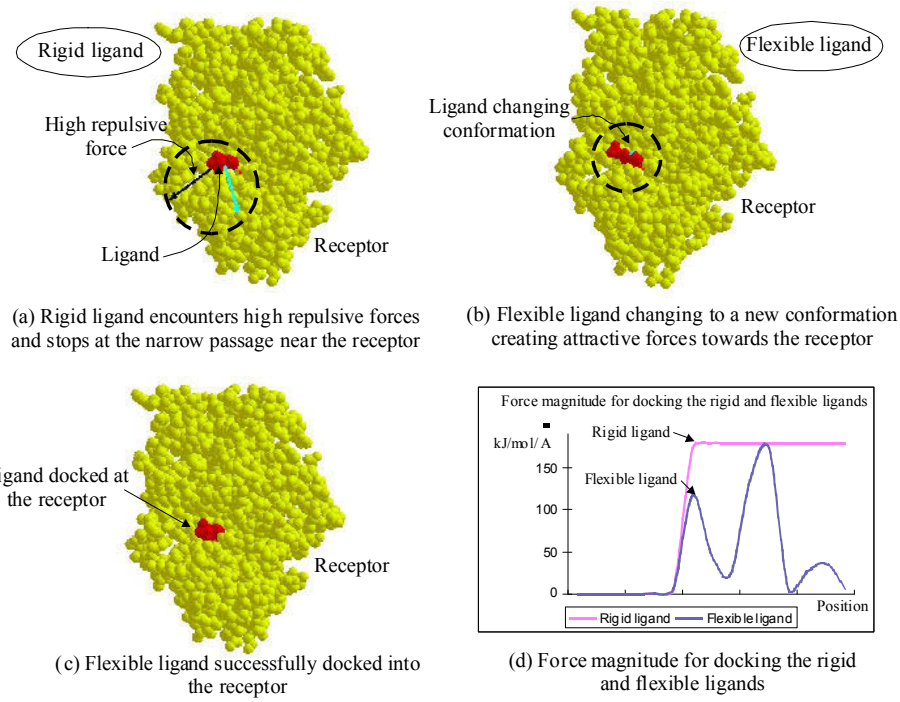


Fig. 6. Results of the proposed adaptive local search algorithm for real-time flexible ligand conformational search.

Figure 7 shows the difference between docking/assembling a rigid and a flexible Tyrosine ligand into the Tyrosyl-tRNA synthetase receptor. When a rigid ligand is used, it is observed that a collision exists and the ligand encounters high repulsive forces as shown in Figure 7(a). On the other hand, when a flexible ligand (9 DOFs) is found by the developed searching algorithm, the ligand changes its flexible conformation and an attractive force towards the receptor is created as shown in Figure 7(b). The attractive force generated by the ligand's new conformation makes the docking/assembly to the receptor possible. Figure 7(c) shows the ligand docked/assemble to the receptor.
Figure 7(d) shows the different force responses of using the rigid ligand and the flexible ligand. As shown in Figure 7(d), the flexible ligand changes into a new conformation with significantly less resistant force and successfully docks/assembles into the receptor. Through the interactive haptic system, the user experiences the change of molecular responding forces and torques from repulsive to attractive while performing the design and simulation of molecular docking and assembly.



(a) Rigid ligand encounters high repulsive forces and stops at the narrow passage near the receptor

(b) Flexible ligand changing to a new conformation creating attractive forces towards the receptor

(c) Flexible ligand successfully docked into the receptor

(d) Force magnitude for docking the rigid and flexible ligands

Fig. 7. Results for docking/assembling rigid and flexible ligands to a receptor.

6. CONCLUSIONS

In this paper, a new method of computer-aided molecular design (CAMD) with the haptic force-torque rendering has been presented. An adaptive local search and energy minimization method has been developed for molecular docking and assembly. The molecular flexibility in terms of its torsional angles is considered to find new collision-free and low-energy molecular conformations that will facilitate the docking and assembly of molecules. The presented techniques can be used for the computer-aided molecular design (CAMD) to provide researchers a real-time interactive computer-aided design for the visualization, manipulation and assembly of molecules in a virtual environment.

7. ACKNOWLEDGMENTS

This work was partially supported by the National Science Foundation (NSF) Grant (DMI-0300297) and the Army Research Office (Grant #W911NF-04-D-0003) to North Carolina State University. Their support is greatly appreciated. The author would also like to thank Dr. Donald Brenner of the Materials Science and Engineering Department at NCSU and Dr. Sanjay Sarma of Mechanical Engineering Department at MIT for their helpful discussion and suggestions.

8. REFERENCES

[1] Bayazit, O. B., Song, G., and Amato, N. M., Ligand binding with OBPRM and user input, *Proceedings of the IEEE International Conference on Robotics & Automation*, Seoul, Korea, May 21-26, 2001, pp 954-959.

[2] Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. and Bourne, P. E., The Protein Data Bank, *Nucleic Acids Research*, Vol. 28, 2000, pp 235-242.

[3] Dubey, A., Sharma, G., Mavroidis, C., Tomassone, S. M., Nikitczuk, K., and Yarmush, M. L., Dynamics and kinematics of viral protein linear nano-actuators for bio-nano robotic systems, *Proceedings of the IEEE International Conference of Robotics and Automation*, New Orleans, LA, April 26-May 1, 2004, pp 1628-1633.

[4] Lai-Yuen, S. K., and Lee, Y.-S., NanoDAS: A nano-scale docking and assembly simulator for computer-aided molecular design (CAMD), (in revision), Submitted to *Computer-Aided Design*, 2004.

[5] Lai-Yuen, S. K., and Lee, Y.-S., Computer-aided molecular design (CAMD) with force-torque feedback, *Proceedings of the Ninth International Conference on Computer Aided Design and Computer Graphics*, Hong Kong, China, December 7-10, 2005, pp. 199-204.

[6] Lai-Yuen, S. K., Lee, Y.-S., and Zhu, W., Haptic force-torque rendering and energy minimization for computer-aided molecular design (CAMD) and assembly simulation, (in revision) Submitted to the *Computer-Aided Design*, 2005.

[7] Lai-Yuen, S. K., and Lee, Y.-S., Haptic-based automatic molecular docking and assembly search system for computer-aided molecular design (CAMD), To appear in *Proceedings of the IEEE Virtual Reality Conference*, Alexandria, VA, March 25-29, 2006.

[8] Lai-Yuen, S. K., and Lee, Y.-S., Computer-aided design and simulation for nano-scale assembly, To appear in the *Transactions of the North American Manufacturing Research Conference (NAMRC)*, Milwaukee, WI, May 23-26, 2006.

[9] LaValle, S. M., Finn, P. W., Kavraki, L. E., and Latombe, J.-C., A randomized kinematics-based approach to pharmacophore-constrained conformational search and database screening, *Journal of Computational Chemistry*, Vol. 21, No. 9, 2000, pp 731-747.

[10] Lee, Y.-G., and Lyons, K. W., Smoothing haptic interaction using molecular force calculations, *Computer-Aided Design*, Vol. 36, No. 1, 2004, pp 75-90.

[11] Meng, E. C., Shoichet, B. K., and Kuntz, I. D., Automated docking with grid-based energy evaluation, *Journal of Computational Chemistry*, 13(4), 1992, pp. 505-524.

[12] Nagata, H., Mizushima, H., and Tanaka, H., Concept and prototype of protein-ligand docking simulator with force feedback technology, *Bioinformatics*, Vol. 18, No. 1, 2002, pp 140-146.

[13] Ren, Y., Lai-Yuen, S. K., and Lee, Y.-S., Virtual prototyping and manufacturing planning by using tri-dexel models and haptic force feedback, *Virtual and Physical Prototyping*, Vol. 1, No. 1, pp 3-18.

[14] Sharma, G., Mavroidis, C. and Ferreira, A., Virtual reality and haptics in nano- and bionanotechnology, *Handbook of Theoretical and Computational Nanotechnology*, Edited by M. Rieth and W. Schommers, American Scientific Publishers, 2005.

[15] Zhang, M., and Kavraki, L. E., A new method for fast and accurate derivation of molecular conformations, *Journal of Chemical Information and Computing Sciences*, Vol. 42, 2002, pp 64-70.

[16] Zhu, W., and Lee, Y-S., A visibility sphere marching algorithm of constructing polyhedral models for haptic sculpting and product prototyping," *Robotics and Computer Integrated Manufacturing,* Vol. 21, No. 1, 2005, pp. 19-36.